AD 293 677

UNCLASSIFIED

NOTICE: When government or other drawings, specifications or other data are used for any purpose other than in connection with a definitely related government procurement operation, the U. S. Government thereby incurs no responsibility, nor any obligation whatsoever; and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data is not to be regarded by implication or otherwise as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use or sell any patented invention that may in any way be related thereto.

The RAND Corporation
SANTA MONICA · CALIFORNIA

55-2441 C.1

STUDIES ON BOTTLENECK PROBLEMS IN PRODUCTION PROCESSES

Richard Bellman
R. Sherman Lehman

P-492

25 February 1954

The RAND Corporation
1700 MAIN ST. • SANTA MONICA • CALIFORNIA

53 pp

P R E F A C E

"Studies on Bottleneck Problems in Production Processes" consists of two papers are given, their purpose being to study in detail some representative problems which occur in the study of inter-industry processes. The general techniques will be those of the theory of dynamic programming. A brief description of these techniques appears in [1]; a longer expository account, together with ~~an application~~, appears in [2].

The first paper, by R. Sherman Lehman, considers a simple mathematical model of the interaction of two interdependent industries engaged in the production of one item. For convenience, ~~we shall call these~~ the auto and steel industries are considered.

Utilizing a highly lumped model, we shall assume that at any time t the state of the industries is completely specified by the following quantities:

to pg iii

$$
(1)\qquad
\begin{aligned}
x_1(t) &= \text{auto stockpile at time } t\\
x_2(t) &= \text{auto capacity at time } t\\
x_3(t) &= \text{steel stockpile at time } t\\
x_4(t) &= \text{steel capacity at time } t
\end{aligned}
$$

Taking t to be a continuous variable, at each instant we must determine rates of allocation of the steel stockpile towards three distinct objectives:

(2)
- (a) Production of autos
- (b) Building of auto factories, i.e., increase of auto capacity
- (c) Building of steel factories; i.e., increase of steel capacity

The last two of these three objectives are to be sublimated to the over-all objective of maximizing the total number of autos produced over a time-period T, which is to say, the quantity $x_1(T)$.

The measures of stockpile and capacity are chosen so that one unit of capacity, either auto or steel, is required for the production of one unit of stockpile in unit time. We assume that b_1 units of steel are required to make one unit of autos, b_2 units of steel to increase auto capacity by one unit, and b_4 units of steel to increase steel capacity by one unit. However, we shall assume that no steel is required to produce additional steel.

Finally, we shall assume that there are no time-lags between allocation and production, an assumption of some magnitude.

Let

(3)
- (a) $z_1(t)$ = rate of production of autos
- (b) $z_2(t)$ = rate of increase of auto capacity
- (c) $z_3(t)$ = rate of production of steel
- (d) $z_4(t)$ = rate of increase of steel capacity

We obtain the following system of equations:

$$\frac{dx_1}{dt} = z_1(t), \quad x_1(0) = c_1,$$

$$(4) \qquad \frac{dx_2}{dt} = z_2(t), \quad x_2(0) = c_1,$$

$$\frac{dx_3}{dt} = z_3(t) - b_1 z_1(t) - b_2 z_2(t) - b_4 z_4(t), \; x_3(0) = c_3,$$

$$\frac{dx_4}{dt} = z_4(t), \quad x_4(0) = c_4,$$

together with the linear constraints

$$(5) \qquad \begin{array}{ll} \text{(a)} & z_1(t) \leq x_2(t) \\ \text{(b)} & z_3(t) \leq x_4(t) \\ \text{(c)} & z_i(t) \geq 0, \quad i=1,2,3,4 \\ \text{(d)} & x_3(t) \geq 0. \end{array}$$

The first two of these constraints arise from the capacity limitations. The third is a statement that the rates of production must be non-negative, and the fourth that the steel stockpile must be non-negative.

The problem is now to determine the $z_i(t)$, satisfying the restrictions of (5) which maximize $x_1(T)$.

The second paper, by Richard Bellman and R. Sherman Lehman, treats the problem of maximizing $x_2(T)$ given the equations

dx sub 2/dt

$$(6) \qquad \frac{dx_2}{dt} = a_2 z_2(t) - z_3(t), \qquad x_2(0) < c_2,$$ and

dx sub 3/dt

$$\frac{dx_3}{dt} = b_3 \gamma_3 z_3(t) - \gamma_2 z_2(t), \quad x_3(0) = c_3,$$

* sub 2   ** sub 3   *** gamma

and the constraints

$$\text{(7)}\quad \begin{aligned} &\text{(a)}\quad z_2,\ z_3 \geq 0 \\ &\text{(b)}\quad z_2 + z_3 \leq x_2 \\ &\text{(c)}\quad \gamma_2 z_2 + \gamma_3 z_3 \leq x_3 \end{aligned}$$

greater than or equal to 0,
less than or equal to x sub 2,
less than or equal to x sub 3,

where a_2, b_2, γ_2, γ_3 are positive constants.

This problem is derived from the same type of mathematical model as the first.

Both problems require for their treatment an intimate intermingling of general theory and detailed analysis. It seems fairly certain that it is not altogether the fault of our methods that the analysis is so complicated, but that these are actually complicated problems with nonobvious solutions.

* sub 2
** sub 3
*** gamma

STUDIES ON BOTTLENECK PROBLEMS IN PRODUCTION PROCESSES

STUDY I:

A SIMPLE TWO-INDUSTRY PROBLEM INVOLVING MULTI-STAGE PRODUCTION PROCESSES

by

R. Sherman Lehman

§1. Introduction

In the Preface to these Studies, we have formulated in mathematical terms the problem of utilizing the steel and auto industries so as to maximize auto production. Let us then continue from equations (4) and (5) of the Preface.

These equations can be combined to provide an equivalent system of integral inequalities:

$$
\begin{aligned}
& z_1 \le x_2: \quad z_1(t) - \int_0^t z_2(s)ds \le c_2 \\
(1.1) \qquad & 0 \le x_3: \quad \int_0^t (-z_3(s) + b_1 z_1(s) + b_2 z_2(s) + b_4 z_4(s))ds \le c_3 \\
& z_3 \le x_4: \quad z_3(t) - \int_0^t z_4(s)ds \le c_4
\end{aligned}
$$

Our problem is a special case of the following more general problem. Let Z be the set of all vector functions $z(t)$ which satisfy the conditions*

$$
(1.2) \qquad \begin{aligned}
& z(t) \ge 0 \\
& Bz(t) + \int_0^t Cz(s)ds \le c
\end{aligned}
$$

* If u and v are vectors, by the notation $u \le v$ we mean that each component of $u \le$ the corresponding component of v, cf. [2].

where B and C are matrices, and c is a constant vector. The problem is to find a vector function z(t) in Z which maximizes

$$(1.3) \qquad \int_0^T (z(t),\alpha)dt,$$

the integral of the inner product of z with a constant vector α.

In [2], it was shown that there is a dual problem which provides a sufficient condition that a z(t) belonging to Z be a maximizing vector function, or in other words, that a feasible solution be optimal. Let W be the set of vector functions w(t) for which

$$(1.4) \qquad \begin{aligned} & w(t) \geq 0 \\ & B'w(t) + C' \int_t^T w(s)ds \geq \alpha \end{aligned}$$

where B' and C' are the transposes of B and C. The dual problem is to find $\underset{w \in W}{\mathrm{Min}} \int_0^T (w(t),c)dt$. It can be proved readily that for

$$(1.5) \qquad \int_0^T (z(t),\alpha)dt \leq \int_0^T (w(t),c)dt,$$

and that if we find two vector functions for which (1.5) holds with equality, they must yield the maximum and minimum, respectively, for the two problems (see [2]). Two such vector functions for which equality holds will be said to be paired with each other. Thus, a sufficient condition that a z belonging to Z be optimal is that it can be paired with some w in W.

For the auto-steel problem formulated above we have

$$(1.6)\qquad B = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \qquad C = \begin{pmatrix} 0 & -1 & 0 & 0 \\ b_1 & b_2 & -1 & b_4 \\ 0 & 0 & 0 & -1 \end{pmatrix} \qquad \alpha = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

The dual system of inequalities is therefore

$$(1.7)\qquad \begin{aligned} \ell_1 &\equiv w_2(t) + b_1 \int_t^T w_3(s)ds - 1 \geq 0 \\ \ell_2 &\equiv -\int_t^T w_2(s)ds + b_2 \int_t^T w_3(s)ds \geq 0 \\ \ell_3 &\equiv w_4(t) - \int_t^T w_3(s)\, ds \geq 0 \\ \ell_4 &\equiv b_4 \int_t^T w_3(s)ds - \int_t^T w_4(s)ds \geq 0. \end{aligned}$$

We have chosen to call the components of w, w_2, w_3, and w_4 in order to keep the connection with the inequalities $z_1 \leq x_2$, $0 \leq x_3$, $z_3 \leq x_4$ clear.

The optimality conditions, i.e., the conditions that (1.5) hold with equality, are:

$$(1.8)\qquad \begin{aligned} &\text{If } z_i(t) > 0, \text{ then } \ell_i(t) = 0, \quad (i=1,2,3,4) \\ &\text{If } z_1(t) < x_2(t), \text{ then } w_2(t) = 0 \\ &\text{If } 0 < x_3(t), \text{ then } w_3(t) = 0 \\ &\text{If } z_3(t) < x_4(t), \text{ then } w_4(t) = 0 \end{aligned}$$

The following are equivalent to the optimality conditions:

If $\ell_i(t) > 0$, then $z_i(t) = 0$, $(i=1,2,3,4)$

If $w_2(t) > 0$, then $z_1(t) = x_2(t)$

If $w_3(t) > 0$, then $0 = x_3(t)$

If $w_4(t) > 0$, then $z_3(t) = x_4(t)$

§2. Delta-functions

Before we proceed to determine the solution, let us discuss the use that we will make of delta functions. It can easily happen that the general problems discussed above have no solutions if the sets Z and W are composed only of vectors having components which are measurable functions. In fact, as we shall later see, this is the usual case in the auto-steel problem. This difficulty can be evaded by enlarging the sets Z and W so that they contain vector "functions" whose components are sums of measurable functions and "delta functions." In these enlarged classes the problems have solutions. By a delta function concentrated at t_0 with weight ω, which se denote by $\omega\delta(t-t_0)$, we mean an improper function such that

$$\int_0^t \delta(s-t_0)\phi(s)ds = \begin{matrix} 0 \text{ if } t < t_0 \\ \omega\phi(t_0) \text{ if } t > t_0 \end{matrix}$$

for every function ϕ continuous at t_0. (For $t = t_0$ the integral is undefined except when $\phi(t_0) = 0$, in which case it is defined to be 0.)

The use of delta functions can be justified rigorously either by the alternative use of Stieltjes integrals, or by regarding the delta functions as obtained by completing the space of measurable functions by a process similar to that used in obtaining the real numbers from the rationals.*

The optimality conditions remain the same even when Z and W are enlarged in the above way. We observe that there is no harm in the violation of the optimality conditions at isolated points or even in sets of measure zero when only measurable functions are allowed as components of z and w. But, when one of the vectors, w, for example, has a component w_1 which is a delta function at the point t_o, then for a z to be paired with w, the corresponding optimality conditions must be satisfied at the point t_o.

We shall find that we never have to use delta functions concentrated at any point other than 0 to obtain an optimal z. Intuitively, this means that radical changes are not necessary except at the beginning.

§3. The Solution

The procedure that we use will be to construct a number of w-solutions which we can pair with z's belonging to Z and hence obtain solutions of our problem. The chief difficulty occurs in constructing w-solutions with suitable properties. In this we are guided by a combination of guessing and observation of properties that an optimal z should have.

---

* Cf. G. Temple, "Theoreis and Applications of Generalized Functions," Jour. Lond. Math. Soc., 28 (1953), pp. 134-148.

First of all, it is clear that we should always have $z_3 = x_4$. To produce too much steel is not harmful. This tells us that we should have $\ell_3(t) = 0$ for all t; i.e., $w_4(t) = \int_t^T w_3(s)ds$. The remaining inequalities of (1.7) then become

$$
\begin{aligned}
&\ell_1 \equiv w_2(t) + b_1\dot{w}_4(t) - 1 \geq 0 \\
(3.1)\qquad &\ell_2 \equiv b_2 w_4(t) - \int_t^T w_2(s)ds \geq 0 \\
&\ell_4 \equiv b_4 w_4(t) - \int_t^T w_4(s)ds \geq 0
\end{aligned}
$$

Shortly before T it is clear that we should be producing x_1 since that is the quantity we wish to maximize at time T. Hence, we will have $z_1 > 0$, which implies that $\ell_1 = 0$. This alone will not give us sufficient information to determine w_2 and w_4.

We first construct a w solution, which we shall call the basic w-solution, with the property that $\ell_2 = 0$ near the end. This means that we must have $w_4(T) = 0$. Then by (2.1) we have

$$
\begin{aligned}
&w_4(t) = \frac{1}{b_1}\left(1-e^{-b_1(T-t)/b_2}\right) \\
(3.2)\qquad &w_2(t) = e^{-b_1(T-t)/b_2}
\end{aligned}
$$

We see that w_2, w_3, and w_4 all remain positive as t decreases. We must check to see whether the inequality $\ell_4 \geq 0$ is satisfied. With the above choice of w we have

$$\ell_4 = \frac{b_4}{b_1}\left(1-e^{-b_1(T-t)/b_2}\right) - \frac{(T-t)}{b_1} + \frac{b_2}{b_1^2}\left(1-e^{-b_1(T-t)/b_2}\right) \tag{3.3}$$

The quantity on the right side of this equation is positive for $T-t$ small but is negative when $T-t$ is large. Let t_o be the value of t for which the right side becomes zero. Then $T-t_o$ is the solution of the equation

$$T-t_o = \left(b_4 + \frac{b_2}{b_1}\right)\left(1-e^{-b_1(T-t_o)/b_2}\right) \tag{3.4}$$

Thus we see that at t_o we must abandon one of the equations $\ell_1 = 0$ and $\ell_2 = 0$. Let us try to choose w so that $\ell_1 = 0$ and $\ell_4 = 0$ before t_o. We have

$$\begin{aligned} w_4(t) &= w_4(t_o)e^{(t_o-t)/b_4} \\ w_2(t) &= 1 - b_1 w_4(t_o)e^{(t_o-t)/b_4} \end{aligned} \tag{3.5}$$

To verify that $\ell_2 \geq 0$ we compute its derivative. We find

$$\frac{d\ell_2}{dt} = b_2\frac{dw_4}{dt} + w_2 = 1 - \left(b_1 + \frac{b_2}{b_4}\right)w_4(t_o)e^{(t_o-t)/b_4} \tag{3.6}$$

A condition sufficient to insure that $\ell_2 \geq 0$ is that $\frac{d\ell_2}{dt} \leq 0$ for all $t \leq t_o$. This will hold if

$$w_4(t_o) \geq \frac{b_4}{b_2+b_1 b_4} \tag{3.7}$$

which by (3.2) and (3.4) is equivalent to $T-t_o \geq b_4$. This last inequality can be checked by putting b_4 in place of $T-t$ in (3.3) and verifying that the quantity thus obtained is positive. We have

$$(3.8) \qquad \frac{1}{b_1}\left[(b_4 + \frac{b_2}{b_1})(1-e^{-b_1 b_4/b_2}) - b_4\right]$$

$$= \frac{b_2}{b_1^2} e^{-b_1 b_4/b_2}\left[e^{b_1 b_4/b_2} - (1 + \frac{b_1 b_4}{b_2})\right] > 0.$$

Hence $\ell_2 \geq 0$ for all $t \leq t_o$ with the above choice of w.

We also see from (3.5) that w_4 and w_3 remain positive. Thus (3.5) will give a satisfactory choice of w until w_2 becomes zero. Let t_1 be the value of t when this happens. Then, by (3.2) and (3.4) we have

$$(3.9) \qquad e^{(t_o-t_1)/b_4} = \frac{b_4 + b_2/b_1}{(T-t_o)} \quad .$$

Before t_1 let us see whether we can choose $w_2 = 0$ and have $\ell_4 = 0$. We see that $w_4 > 0$ and $w_3 > 0$. We have $\frac{d\ell_2}{dt} = b_2\frac{dw_4}{dt} < 0$ so that $\ell_2 > 0$, and $\frac{d\ell_1}{dt} = b_1\frac{dw_4}{dt} < 0$ so that $\ell_1 > 0$. Hence this choice of w will be valid for all $t \leq t_1$.

Our basic solution is summarized in the following table. This table also lists the properties that a z paired with this w solution must have. Any z with these properties gives a policy which, if feasible (i.e., satisfies the z constraints), is optimal.

(3.10)

$t < t_1$		$t_1 < t < t_o$		$t_o < t < T$	
$\ell_1 > 0$	$z_1 = 0$	$\ell_1 = 0$		$\ell_1 = 0$	
$\ell_2 > 0$	$z_2 = 0$	$\ell_2 > 0$	$z_2 = 0$	$\ell_2 = 0$	
$\ell_3 = 0$		$\ell_3 = 0$		$\ell_3 = 0$	
$\ell_4 = 0$		$\ell_4 = 0$		$\ell_4 > 0$	$z_4 = 0$
$w_2 = 0$		$w_2 > 0$	$z_1 = x_2$	$w_2 > 0$	$z_1 = x_2$
$w_3 > 0$	$x_3 = 0$	$w_3 > 0$	$x_3 = 0$	$w_3 > 0$	$x_3 = 0$
$w_4 > 0$	$z_3 = x_4$	$w_4 > 0$	$z_3 = x_4$	$w_4 > 0$	$z_3 = x_4$

Fig. 1

Let us see how this table can be used to obtain a partial solution of the auto-steel problem. For the moment let us assume $c_3 = 0$. For $t < t_1$ we must have

$$(3.11)\qquad z_1 = 0, \quad z_2 = 0, \quad z_3 = x_4, \quad z_4 = x_4/b_4.$$

For $t_1 < t < t_o$ we must choose

$$(3.12)\qquad z_1 = x_2, \quad z_2 = 0, \quad z_3 = x_4, \quad z_4 = \frac{x_4 - b_1 x_2}{b_4}$$

This can be done if and only if $x_4(t_1) - b_1 x_2(t_1) \geq 0$. Let us assume that this inequality is satisfied. Then for $t_o < t < T$ we must have

$$(3.13)\qquad z_1 = x_2, \quad z_2 = \frac{x_4 - b_1 x_2}{b_2}, \quad z_3 = x_4, \quad z_4 = 0$$

which is possible provided $x_4(t_1) - b_1x_2(t_1) \geq 0$. Thus we see that for certain initial conditions we can obtain the optimal solution.

§4. The Modified w Solution.

As already has been noted, we run into trouble if $x_4(t_1) - b_1x_2(t_1) < 0$. To handle this case we consider a modification of the basic w solution of Fig. 1 above. Let u_o be in the interval $[t_1, T]$. For each such u_o we define a solution as follows:

For $u_o < t < T$ we let $w(t)$ be the same as in the basic solution. For $t < u_o$ we choose $w_2(t) = 0$. For $t < u_o$ but near u_o we choose $w_4(t) = 1/b_1$ so that $\ell_1 = 0$. This choice will keep $\ell_4 > 0$ for a while before u_o. We define u_1 to be the point where ℓ_4 becomes 0 with this choice of w. For $t < u_1$ we choose w so that $\ell_4 = 0$. It is easily seen that this choice makes $\ell_1 > 0$, $\ell_2 > 0$, $w_3 > 0$ and $w_4 > 0$ for all $t < u_1$. Hence, in this way we obtain a w solution for each u_o in the interval $[t_1, T]$. We observe that for $u_o = t_1$, $u_1 = t_1$ and this solution is identical with our basic solution of Fig. 1. Note that u_1 depends continuously on u_o. Since for $u_o = T$, $u_1 = T - b_4$, there is a w solution for each u_1 in the interval $[t_1, T-b_4]$.

These w solutions together with the properties of the corresponding z solutions, are summarized in the following table:

(4.1)

$t < u_1$		$u_1 < t < u_o$		$u_o < t < T$	$t_1 < t < t_o$	$u_o < t < T$	$t_o < t < T$
$\ell_1 > 0$	$z_1 = 0$	$\ell_1 = 0$		$\ell_1 = 0$		$\ell_1 = 0$	
$\ell_2 > 0$	$z_2 = 0$	$\ell_2 > 0$	$z_2 = 0$	$\ell_2 > 0$	$z_2 = 0$	$\ell_2 = 0$	
$\ell_3 = 0$		$\ell_3 = 0$		$\ell_3 = 0$		$\ell_3 = 0$	
$\ell_4 = 0$		$\ell_4 > 0$	$z_4 = 0$	$\ell_4 = 0$		$\ell_4 > 0$	$z_4 = 0$
$w_2 = 0$		$w_2 = 0$		$w_2 > 0$	$z_1 = x_2$	$w_2 > 0$	$z_1 = x_2$
$w_3 > 0$	$x_3 = 0$	$w_3 = 0$		$w_3 > 0$	$x_3 = 0$	$w_3 > 0$	$x_3 = 0$
$w_4 > 0$	$z_3 = x_4$	$w_4 > 0$	$z_3 = x_4$	$w_4 > 0$	$z_3 = x_4$	$w_4 > 0$	$z_3 = x_4$
Since w_3 is a delta function at u_o, we must have $x_3(u_o) = 0$.							

Fig. 2

Note that if $u_o > t_o$, then there is no t satisfying the conditions of the third column; and if $u_o = T$ then there is no t satisfying the conditions of the last column either.

§5. The Equilibrium Solution

A policy which seems plausible in some instances is the following: Make an initial adjustment to bring x_3 down to zero in such a way that after the adjustment $x_4 = b_1 x_2$. If this is done, after the initial adjustment no increase in capacities is necessary and all available steel can be used for auto production. Such a policy would require for the w paird with it that $\ell_2(0) = 0$ and $\ell_4(0) = 0$, because in general both z_2 and z_4 will have to be delta functions. We shall construct a w solution with this property.

First, we note that our basic w solution has this property when T is such that $t_o = 0$. This suggests that we try to choose

$$w_4(t) = \alpha e^{-b_1(T-t)/b_2} + \beta \tag{5.1}$$

where α and β are constants. If w_2 is chosen so that $\ell_1 = 0$, the inequalities (3.1) become

$$\begin{aligned} \ell_2 &\equiv b_2 w_4(t) - (T-t) + b_1 \int_t^T w_4(s)ds \geq 0 \\ \ell_4 &\equiv b_4 w_4(t) - \int_t^T w_4(s)ds \geq 0. \end{aligned} \tag{5.2}$$

If $\ell_2(0) = \ell_4(0) = 0$, then

$$w_4(0) = \frac{T}{b_2 + b_1 b_4} \tag{5.3}$$

We set $E = e^{-b_1 T/b_2}$ and from (5.1) – (5.3) derive the following two equations for α and β:

$$\begin{aligned} & b_2 \alpha + (b_2 + b_1 T)\beta = T \\ & (b_2 + b_1 b_4)E\alpha + (b_2 + b_1 b_4)\beta = T. \end{aligned} \tag{5.4}$$

A solution of these equations will give a w for which $\ell_2(0) = \ell_4(0) = 0$. We have

$$\alpha = \frac{T}{\Delta} \begin{vmatrix} 1 & b_2 + b_1 T \\ 1 & b_2 + b_1 b_4 \end{vmatrix} = \frac{T}{\Delta} \left[b_1 (b_4 - T) \right] \tag{5.5}$$

$$\beta = \frac{T}{\Delta} \begin{vmatrix} b_2 & 1 \\ (b_2 + b_1 b_4)E & 1 \end{vmatrix} = \frac{T}{\Delta} \left[b_2(1-E) - b_1 b_4 E \right]$$

where

$$(5.6)\qquad \Delta = \begin{vmatrix} b_2 & b_2 + b_1 T \\ (b_2 + b_1 b_4)E & b_2 + b_1 b_4 \end{vmatrix} = (b_2 + b_1 b_4)(b_2 - b_2 E - b_1 ET)$$

Also

$$(5.7)\qquad = (b_2 + b_1 b_4)E\ (b_2 e^{b_1 T/b_2} - b_2 - b_1 T)$$
$$(b_2 + b_1 b_4)E\ (b_2 + b_1 T - b_2 - b_1 T) = 0$$

Now let us assume that $T - t_o \geq T \geq b_4$. Then from (5.5) we see that $\alpha \leq 0$. Let us check to be sure that for the w we have defined $w_2(t)$, $w_3(t)$, and $w_4(t)$ are non-negative for $0 \leq t \leq T$. This is equivalent to verifying that $0 \leq w_4(t) \leq 1/b_1$ and $\frac{dw_4}{dt} \leq 0$. We have $\frac{dw_4}{dt} = \alpha \frac{b_1}{b_2} e^{-b_1(T-t)b_2} \leq 0$. Hence it will be sufficient to check that $w_4(T) \geq 0$ and $w_4(0) \leq 1/b_1$. Since $T - t_o \geq T$, we conclude from (3.4) and 5.3) that

$$(5.8)\qquad w_4(0) = \frac{T}{b_2+b_1 b_4} \leq \frac{T-t_o}{b_2+b_1 b_4} < \frac{b_4 + b_2/b_1}{b_2+b_1 b_4} = 1/b_1$$

and

$$(5.9)\qquad w_4(T) = \alpha + \beta = \frac{T}{\Delta} b_1 \left[(b_4 + b_2/b_1)(1-E) - T\right] \geq 0$$

We also must check that for $0 \leq t \leq T$, $\ell_2 \geq 0$ and $\ell_4 \geq 0$.

Since

$$(5.10)\qquad \frac{d\ell_2}{dt} = b_2 \frac{dw_4}{dt} + 1 - b_1 w_4(t) = 1 - b_1\beta$$

and $\ell_2(T) = b_2 w_4(T) \geq 0$, we have $\ell_2 \geq 0$ for all t in $[0,T]$. Similarly, we know that $\ell_4(T) = b_4 w_4(T) \geq 0$. Hence, if we show that $\frac{d^2\ell_4}{dt^2} \leq 0$, we will have proved that $\ell_4 \geq 0$ for all t in $[0,T]$. We have

$$(5.11)\qquad \frac{d^2\ell_4}{dt^2} = a\left(b_4 + \frac{b_2}{b_1}\right)\left(\frac{b_1}{b_2}\right)^2 \leq 0.$$

This completes the proof that the w which we have defined is a solution. Its properties, together with those a z paired with it must have, are summarized in the following table:

(5.12)

$t = 0$	$0 < t < T$	
	$\ell_1 = 0$	
$\ell_2 = 0$	$\ell_2 > 0$	$z_2 = 0$
	$\ell_3 = 0$	
$\ell_4 = 0$	$\ell_4 > 0$	$z_4 = 0$
	$w_2 > 0$	$z_1 = x_2$
	$w_3 > 0$	$x_3 = 0$
	$w_4 > 0$	$x_3 = x_4$

Fig. 3

Note: This solution is valid only for $T \geq b_4$.

§6. A Short-time w Solution

The w solution which we construct next will be useful in finding the solution of our maximum problem when the total time is short, $T < b_4$. This solution differs from those already constructed in that it allows x_3 to be positive and z_2 to be a delta function concentrated at 0.

For $0 \leq t \leq T$ let $w_4(t) = \gamma$, $w_2(t) = 1 - b_1\gamma$ where $0 < \gamma < 1/b_1$. Then $\ell_1(t) = 0$, $\ell_4(t) > 0$ for $0 \leq t < T$. Also

$$\ell_2(t) = b_2\gamma - (T-t)(1-b_1\gamma) = [b_2+b_1(T-t)]\gamma - (T-t) \tag{6.1}$$

Now, if we choose $\gamma = \frac{T}{b_2+b_1T}$ then $\ell_2(0) = 0$ and $\ell_2(t) > 0$ for $t > 0$. Thus we obtain a solution of the system of inequalities (3.1). It is summarized below together with the properties a z paired with it must have.

(6.2)

$t = 0$	$0 < t < T$	
	$\ell_1 = 0$	
$\ell_2 = 0$	$\ell_2 > 0$	$z_2 = 0$
	$\ell_3 = 0$	
	$\ell_4 > 0$	$z_4 = 0$
	$w_2 > 0$	$z_1 = x_2$
	$w_3 = 0$	
	$w_4 > 0$	$z_3 = x_4$
Since w_3 is a delta function at T, $x_3(T) = 0$.		

Fig. 4

Note: This solution is valid for $T < b_4$ only.

§7. Description of Solution and Proof

We now can give the complete solution to the original problem. There are quite a few cases that we must consider separately. The critical values t_o and t_1, which are defined by (3.4) and (3.9), depend on T, but in such a way that for fixed b_1, b_2, and b_4, $T - t_o$ and $T - t_1$ are constants.

Case I: T is large enough so that $t_1 \geq 0$. In this case we choose z_4 to be a delta function concentrated at 0 to bring x_3 down to zero immediately. This means that if the total time is long enough we should not keep any steel in storage but should be using it to build more steel plants. The use of the delta function is permissible because $\ell_4 = 0$ for t near 0. For $0 < t < t_1$ we let

$$(7.1) \qquad z_1 = 0, \quad z_2 + 0, \quad z_3 = x_4, \quad z_4 = x_4/b_4$$

thus keeping x_3 at zero level. At t_1 we must distinguish different subcases:

$$(7.2) \qquad \begin{array}{ll} \text{IA:} & x_4(t_1) - b_1 x_2(t_1) \geq 0 \\ \text{IB:} & x_4(t_1) - b_1 x_2(t_1) < 0 \end{array}$$

In case IA we can produce autos at capacity without running out of steel. Hence we let

$$(7.3) \qquad z_1 = x_2, \quad z_2 = 0, \quad z_3 = x_4, \quad z_4 = \frac{x_4 - b_1 x_2}{b_4}$$

for $t_1 \leq t \leq t_o$; and for $t_o < t \leq T$ we let

$$(7.4) \qquad z_1 = x_2, \quad z_2 = \frac{x_4 - b_1 x_2}{b_2}, \quad z_3 = x_4, \quad z_4 = 0.$$

This solution for Case IA is optimal because it can be paired with our basic w solution of Fig. 1.

In Case IB we do not have enough steel to produce autos at capacity. Hence we continue to produce no autos for $t > t_1$, i.e.,

$$(7.5) \qquad z_1 = 0, \quad z_2 = 0, \quad z_3 = x_4, \quad z_4 = \frac{x_4}{b_4}.$$

We do this until $x_4 - b_1 x_2$ becomes zero or $t = T - b_4$, whichever happens first. If $x_4 - b_1 x_2$ becomes zero at t' then we choose $z_1 = x_2$, $z_2 = 0$, $z_3 = x_4$, $z_4 = 0$ thereafter. This solution is seen to be optimal by pairing it with the w solution of Fig. 2 for which $u_1 = t'$. As we have already remarked there is such a solution no matter what t' is, so long as $t_1 < t' \leq T - b_4$. If, on the other hand, $x_4(T-b_4) - b_1 x_2(T-b_4) < 0$, then for $T - b_4 < t \leq T$ we choose

$$(7.6) \qquad z_1 = \frac{x_4}{b_1}, \; z_2 = 0, \; z_3 = x_4, \; z_4 = 0.$$

This solution can be seen to be optimal by pairing it with the w solution of Fig. 2, for which $u_o = T$, $u_1 = T - b_4$.

<u>Case</u> II: T <u>is such that</u> $t_1 \leq 0 \leq t_o$. As before we choose z_4 to be a delta function concentrated at 0 to bring x_3 down to zero immediately. Thereafter the solution is as before. There are two subcases:

$$(7.7) \qquad \begin{array}{ll} \text{IIA:} & x_4(0) - b_1x_2(0) \geq 0 \\ \text{IIB:} & x_4(0) - b_1x_2(0) < 0 \end{array}$$

In Case IIA we let $z_1 = x_2$, i.e., produce autos at capacity. We use the remaining steel to increase steel capacity before t_o and to increase auto capacity after t_o. That is, for $0 < t < t_o$ we let

$$(7.8) \qquad z_1 = x_2, \quad z_2 = 0, \quad z_3 = x_4, \quad z_4 = \frac{x_4 - b_1x_2}{b_4}$$

and for $t > t_o$ we let

$$(7.9) \qquad z_1 = x_2, \quad z_2 = \frac{x_4 - b_1x_2}{b_4}, \quad z_3 = x_4, \quad z_4 = 0$$

This solution is optimal because it can be paired with our basic solution of Fig. 1.

Case IIB is similar to IB. The same prescription holds, and the solution is paired with one from Fig. 2.

<u>Case</u> III: T <u>is such that</u> $t_o \leq 0 \leq T - b_4$. There are three subcases:

$$
(7.10)\quad
\begin{aligned}
&\text{IIIA:} && c_4 - b_1c_2 \geq b_1\frac{c_3}{b_2}\\
&\text{IIIB:} && c_4 - b_1c_2 < -\frac{c_3}{b_4}\\
&\text{IIIC:} && \frac{-c_3}{b_4} \leq c_4 - b_1c_2 < \frac{b_1c_3}{b_2}
\end{aligned}
$$

In Case IIIA we use our initial stockpile of steel to increase auto capacity, i.e., we let z_2 be a delta function concentrated at 0 bringing x_3 down to zero. Thereafter, we let $z_1 = x_2$ and use any remaining steel to increase auto capacity, i.e.,

$$
(7.11)\qquad z_1 = x_2, \quad z_2 = \frac{x_4 - b_1x_2}{b_2}, \quad z_3 = x_4, \quad z_4 = 0
$$

This solution is optimal because it can be paired with the basic w solution of Fig. 1.

In Case IIIB we find ourselves short on steel capacity. The policy and proof are the same as in Case IB.

In Case IIIC we can make an initial adjustment so that x_3 becomes zero and $x_4 = b_1x_2$. We do this by choosing z_2 and z_4 to be delta functions concentrated at 0. After that we let $z_1 = x_2$, $z_2 = 0$, $z_3 = x_4$, $z_4 = 0$. This solution is optimal because it can be paired with the equilibrium w solution of Fig. 3.

<u>Case</u> IV: $T \leq b_4$. There are three subcases which depend on the initial values:

$$\text{(7.12)} \qquad \begin{aligned} &\text{IVA:} \quad c_4 - b_1c_2 \geq \frac{b_1c_3}{b_2} \\ &\text{IVB:} \quad c_4 - b_1c_2 > \frac{-c_3}{T} \\ &\text{IVC:} \quad \frac{-c_3}{b_4} \leq c_4 - b_1c_2 < \frac{b_1}{b_2}c_3 \end{aligned}$$

In Case IVA the solution and proof are the same as in Case IIIA.

In Case IVB we choose $z_2 = 0$ and $z_4 = 0$ for all t. As always we let $z_3 = x_4$. We choose z_1 in any way such that $z_1(t) \leq x_2(t)$ and $x_3(T) = 0$. Thus, in this case the solution is not unique. Any solution of this form can be seen to be optimal by pairing it with the w solution of Fig. 2 for which $u_0 = T$.

In case IVC we find ourselves in an intermediate case, unable to follow the policies suggested by IVA and B. In this case we make an initial adjustment of the steel stockpile down to the value c_3', using this steel to increase auto capacity. Thereafter we choose $z_1 = x_2$, $z_2 = 0$, $z_3 = x_4$, and $z_4 = 0$. The value c_3' is determined so that $x_3(T) = 0$. It is found that

$$\text{(7.13)} \qquad c_3' = \frac{b_1c_3 - b_2(c_4 - b_1c_2)}{b_2 + b_1T}$$

has this property. This solution is optimal because it can be paired with the short-time w solution of Fig. 4.

SUMMARY

Initial Adjustments

Cases	I: $t_1 \geq 0$.	II: $t_1 \leq 0 \leq t_0$	III: $t_0 \leq 0 \leq T - b_4$	IV: $T < b_4$
A	Adjust x_3 to 0 by increasing x_4.		Bring x_3 to 0 by increasing x_2 "Build auto capacity".	
B	"Build steel capacity".			No initial adjustments.
C	No Case	No Case	Adjust so $x_3 = 0$, $x_4 = b_1 x_2$, by increasing x_2 and x_4.	Adjust x_3 downward, but not to 0, so that $x_3(T) = 0$. Increase x_2.

After the initial adjustments the optimal policy can be determined by a priority system. Before t_1, building steel capacity, i.e., z_4, has first priority. This continues after t_1 until either $x_4 \geq b_1 x_2$ or $t - b_4$, whichever comes first. When this happens, which may be at t_1, of course, first priority is given to auto production, z_1. This will use up all available steel unless $x_4(t_1) > b_1 x_2(t_1)$. In that case second priority is given to building steel capacity until the time t_0. After t_0 second priority is given to building auto capacity.

STUDY II:
A MATHEMATICAL MODEL OF TWO INTERDEPENDENT PROCESSES

by

Richard Bellman
R. Sherman Lehman

§1. Introduction

In this paper we shall study a simple mathematical model of the interaction of two processes.

Using the fundamental device of the theory of dynamic programming, we shall convert the original problem involving the maximization of an integral subject to integral constraints into that of solving a functional equation which may be converted into a nonlinear partial differential equation.

Although the over-all approach is straightforward, the details of the analysis are not simple and, in general, a very precise argumentation is required.

§2. Formulation of the Problem.

We shall consider the following maximization problem: Given

$$(2.1)\qquad \begin{aligned} \frac{dx_2}{dt} &= a_2 z_2 - z_3, & x_2(0) &= c_2 \\ \frac{dx_3}{dt} &= b_3 \gamma_3 z_3 - \gamma_2 z_2, & x_3(0) &= c_3 \end{aligned}$$

where z_2 and z_3 are functions of t subject to the constraints

$$(2.2)\qquad \begin{aligned} z_2(t) + z_3(t) &\leq x_2(t), & z_2(t) &\geq 0 \\ \gamma_2 z_2(t) + \gamma_3 z_3(t) &\leq x_3(t), & z_3(t) &\geq 0, \end{aligned}$$

determine z_2 and z_3 so as to maximize $x_2(T)$. The constants a_2, b_3, γ_2, γ_3 are all positive as are the initial constants c_2 and c_3.

This problem is equivalent to the functional equation

$$(2.3)\qquad f(c_2, c_3, s+t) = \operatorname*{Max}_{D[0,S]} f(x_2(s), x_3(s), t) \qquad f(c_2, c_3, 0) = c_2.$$

where x_2 and x_3 satisfy (1) and the maximum is over all choice of the decision variables z_2 and z_3 which satisfy (2) on the interval $[0,s]$, and $\operatorname{Max} x_2(T) = f(c_2, c_3, T)$, cf. [1].

A solution of the partial differential equation

$$(4) \quad \frac{\partial f}{\partial t}(c_2,c_3,t) = \underset{D[0]}{\mathrm{Max}} \left[\left(a_2 \frac{\partial f}{\partial c_2} - \gamma_2 \frac{\partial f}{\partial c_3}\right) z_2 + \left(b_3 \gamma_3 \frac{\partial f}{\partial c_3} - \frac{\partial f}{\partial c_2}\right) z_3 \right]$$

where $D[0]$ represents a choice over all $z_i(0)$ satisfying (2), will yield a solution of the functional equation and hence the original maximum problem.

Our procedure is to assume the existence of a solution of (4), a function $f(c_2,c_3,t)$ with continuous partial derivatives in its three arguments and piecewise continuous second derivatives. We assume also that the $z_i(t)$ for this maximum function are piecewise continuous. Using these assumptions we derive the form of the solution.

It is convenient to look at the question in terms of regions in the (r,T) plane where T is the time remaining before the process ends and r is the ratio x_2/x_3 or c_2/c_3. A solution to (4) and also (3) is uniquely determined by the specification of what policy is to be followed for each couple (r,T). For given T the optimal policy depends only on the ratio because of the homogeneity of (1) and (2).

In the quadrant $r \geq 0$, $T \geq 0$ we can find certain regions in which a given policy, for example, $z_2 = x_2$, $z_3 = 0$, is to be followed at $t = 0$. The problem reduces to locating the boundaries of these regions.

§3. B/A <u>is a Monotone Function of</u> r.

First, let us look at the partial differential equation (4). The right side is a linear form in the variables z_2, z_3. Let

$$A = (a_2 \frac{\partial f}{\partial c_2} - \gamma_2 \frac{\partial f}{\partial c_3})$$

$$B = (b_3 \gamma_3 \frac{\partial f}{\partial c_3} - \frac{\partial f}{\partial c_2}).$$

The location of the maximizing point (z_2, z_3) will depend on the ratio B/A. Our first aim is to consider this ratio B/A as a function of r and T. Let $f(c_2, c_3, T) = f_1(c_2/c_3, c_3, T)$. Then

$$\frac{\partial f}{\partial c_2} = \frac{1}{c_3} \frac{\partial f_1}{\partial r}$$

$$\frac{\partial f}{\partial c_3} = - \frac{c_2}{c_3^2} \frac{\partial f_1}{\partial r} + \frac{\partial f_1}{\partial c_3} .$$

Also because of the homogeneity, $f_1(c_2/c_3, ac_3, T) = af_1(c_2/c_3, c_3, T)$; and hence $c_3 \frac{\partial f_1}{\partial c_3} = f_1$. Consequently

$$\frac{\partial f}{\partial c_3} = \frac{1}{c_3}(f_1 - r \frac{\partial f_1}{\partial r})$$

Clearly B/A is a function of r and T alone. For simplicity we consider A and B as functions of r and T by setting $c_3 = 1$, and we drop the subscript on f. We have

$$A = a_2 \frac{\partial f}{\partial r} - \gamma_2 (f - r \frac{\partial f}{\partial r})$$

$$B = b_3 \gamma_3 (f - r \frac{\partial f}{\partial r}) - \frac{\partial f}{\partial r}$$

Considered for fixed T, A is a monotone-decreasing function of r, and B is a monotone-increasing function of r. To prove this

we differentiate with respect to r and find

$$\frac{dA}{dr} = (a_2 + \gamma_2 r)\,\frac{\partial^2 f}{\partial r^2}$$

$$\frac{dB}{dr} = -(1 + b_3\gamma_3 r)\,\frac{\partial^2 f}{\partial r^2}\,.$$

It will be sufficient to prove $\frac{\partial^2 f}{\partial r^2} \leq 0$ because the first derivatives of f are continuous. This will follow from the fact that

$$f(r) \geq \frac{f(r+\Delta) + f(r-\Delta)}{2}\,.$$

To prove this we note that a possible policy for time T with initial values rc_3, c_3 is the mean between the policies which are optimal with initial values $(r+\Delta)c_3$, c_3 and $(r-\Delta)c_3$, c_3 for the same time T. This is true because both the equations and restrictions are linear.

§4. The case $\gamma_3/\gamma_2 < 1$

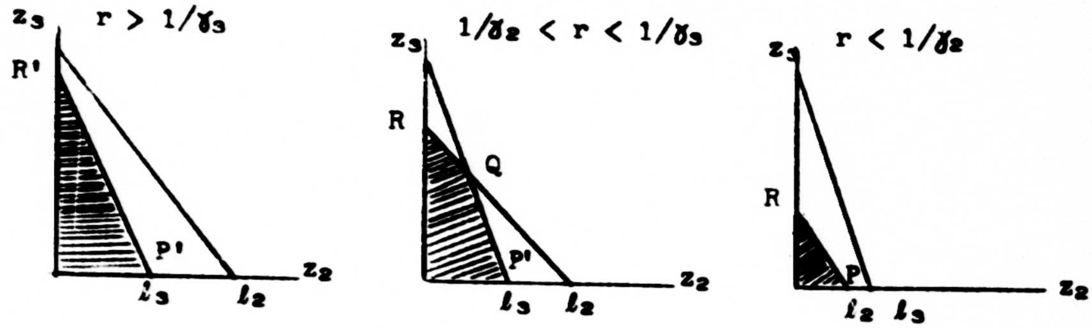
Let us begin by studying the case in which $\gamma_3/\gamma_2 < 1$. If we consider the constraints (2) in the z_2, z_3 plane, we find there are three cases depending on the position of the lines

$$\ell_2:\quad z_2 + z_3 = x_2$$

$$\ell_3:\quad \gamma_2 z_2 + \gamma_3 z_3 = x_3.$$

The line ℓ_2 may be above ℓ_3 so that ℓ_3 is the only essential constraint. The lines may cross so that both are essential. The line ℓ_2 may be below ℓ_3 so that it is the only essential constraint. In

each of these three cases we denote the vertices of the convex set defined by (2) by P, P', Q, R, R'.



For P': $z_2 = x_3/\gamma_2$, $z_3 = 0$

R': $z_2 = 0$, $z_3 = x_3/\gamma_3$

Q: $z_2 = \gamma_3 x_2 - x_3/\gamma_3 - \gamma_2$, $z_3 = x_3 - \gamma_2 x_2/\gamma_3 - \gamma_2$

P: $z_2 = x_2$, $z_3 = 0$

R: $z_2 = 0$, $z_3 = x_2$

The linear form in the right side of (4) will in general be maximized by choosing (z_2, z_3) to be a vertex. However, if the ratio $B/A = 1$ or γ_3/γ_2 the maximizing point (z_2, z_3) is not uniquely determined by the linear form. Consider the trajectory curves for a given solution in the (r, T) space, that is, the curves determined by $\left(\frac{x_3(s)}{x_2(s)}, T-s\right)$ as s runs from 0 to T. If the trajectory never runs along one of the curves $B/A = 1$ or $B/A = \gamma_3/\gamma_2$, then at each point an optimal policy will be to choose (z_2, z_3) to coincide with one of the vertices. A policy of choice of an intermediate point can only be optimal on a curve in the (r, T) plane and not in a region.

Except for this disturbing possibility of choice of an intermediate point, if we knew the location of the curves $B/A = 1$ and $B/A = \gamma_3/\gamma_2$, we would know the optimal policies. They are as follows:

$r > 1/\gamma_3$	$1/\gamma_2 < r < 1/\gamma_3$	$r < 1/\gamma_2$
$B > \frac{\gamma_3}{\gamma_2} A$: R'	$B > A$: R	$B > A$: R
$B < \frac{\gamma_3}{\gamma_2} A$: P'	$A > B > \frac{\gamma_3}{\gamma_2} A$: Q	$B < A$: P
	$B < \frac{\gamma_3}{\gamma_2} A$: P'	

It is clear from the differential equation (1) that for T sufficiently small the optimal policy is to maximize z_2 and take $z_3 = 0$. The favorable effect of taking z_3 positive is of second order while the unfavorable effect of decreasing $\frac{dx_2}{dt}$ is of first order. This means that the optimal policy is P when $r < 1/\gamma_2$ and P' when $r > 1/\gamma_2$.

§5. Region Where P' is Followed.

Knowing the solution for small T, we use the information thus obtained to extend the solution to larger T. First we treat the case in which $r > 1/\gamma_2$ and P' is followed. There

$$\frac{dx_2}{dt} = \frac{a_2}{\gamma_2} x_3$$

$$\frac{dx_3}{dt} = - x_3$$

so that $x_3(t) = c_3e^{-t}$; $x_2(t) = c_2 + \frac{a_2}{\gamma_2} c_3(1-e^{-t})$. Consequently

$$f(c_2,c_3,T) = c_2 + \frac{a_2}{\gamma_2} c_3(1-e^{-T})$$

and

$$A = a_2 e^{-T}$$

$$B = \frac{b_3 \gamma_3 a_2}{\gamma_2}(1 - e^{-T}) - 1.$$

We wish to find when $B = \frac{\gamma_3}{\gamma_2} A$. Then we have

$$e^{-T} = \frac{b_3\gamma_3 a_2 - \gamma_2}{b_3\gamma_3 a_2 + \gamma_3 a_2} = \frac{b_3\gamma_3 a_2 - \gamma_2}{a_2\gamma_3(b_3+1)}.$$

The necessary and sufficient condition that there be such a T is that $\lambda \equiv b_3\gamma_3 a_2 - \gamma_2 > 0$. From (1) one sees that following P' does not decrease r. Define T* by the equation

$$e^{-T^*} = \frac{\lambda}{a_2\gamma_3(b_3+1)}$$

if the right side is positive, and let $T^* = \infty$ otherwise. Then for $r > 1/\gamma_2$, $T < T^*$, we know the solution completely. We note for future reference that this result holds also when $\gamma_3/\gamma_2 \geq 1$.

Let us look at the region $1/\gamma_2 < r < 1/\gamma_3$. Because of the continuity of B/A we know the policy Q is followed for values of T that are slightly larger than T*. We now prove that actually the policy Q is followed in the whole region $1/\gamma_2 < r < 1/\gamma_3$, $T > T^*$.

§6. Q Policy not Preceded by P'.

Let us first establish that in an optimal policy it does not happen that first P' is followed and then Q. For, consider any policy which consists in following P' for $0 < t \leq u - \delta$ and Q for $u - \delta < t \leq u$. We prove that for δ and δ/u sufficiently small there is another policy which is superior having the form Q for $0 < t < \Delta$ and P' for $\Delta < t < u$, where Δ is a positive number.

Let

$$P' = \begin{pmatrix} 0 & a_2/\gamma_2 \\ 0 & -1 \end{pmatrix}$$

$$Q = \frac{1}{(\gamma_3-\gamma_2)} \begin{pmatrix} a_2\gamma_3 + \gamma_2 & -(a_2+1) \\ -\gamma_2\gamma_3(b_3+1) & b_3\gamma_3+\gamma_2 \end{pmatrix}$$

Using vector matrix notation with $x(t) = \begin{pmatrix} x_2(t) \\ x_3(t) \end{pmatrix}$, $c = \begin{pmatrix} c_2 \\ c_3 \end{pmatrix}$, we have $\frac{dx}{dt} = P'x(t)$ when P' is followed, and $\frac{dx}{dt} = Qx(t)$ when Q is followed. Thus for the policies described above we have

$$\text{I:} \quad x(u) = e^{Q\delta}e^{P'(u-\delta)}\cdot c = (I + Q\delta + o(\delta))(I - P'\delta + o(\delta))e^{P'u}\cdot c$$

$$= (I + (Q - P')\delta + o(\delta))e^{P'u}\cdot c$$

$$\text{II:} \quad x(u) = e^{P'(u-\Delta)}e^{Q\Delta} = e^{P'u}(I + (Q - P')\Delta + o(\Delta))\cdot c$$

For δ/u, Δ/u, $u \longrightarrow 0$, the differences between the results at time u obtained by following these policies and $e^{P'u}\cdot c$ which is obtained by following P' for the entire interval are

$$\text{I:} \quad \delta\Big[(Q - P') + (Q - P')P'u + o(u)\Big]\cdot c$$

$$\text{II:} \quad \Delta\Big[(Q - P') + P'(Q-P')u + o(u)\Big]\cdot c.$$

Setting $a_{11} = a_2\gamma_3 + \gamma_2$, $a_{21} = -\gamma_2\gamma_3(b_3+1)$, we find

$$(Q - P') = \frac{1}{(\gamma_3 - \gamma_2)} \begin{pmatrix} a_2\gamma_3 + \gamma_2 & -(a_2\gamma_3 + \gamma_2)/\gamma_2 \\ -\gamma_2\gamma_3(b_3+1) & \gamma_3(b_3+1) \end{pmatrix}$$

$$= \frac{1}{(\gamma_3 - \gamma_2)} \begin{pmatrix} a_{11} & -a_{11}/\gamma_2 \\ a_{21} & -a_{21}/\gamma_2 \end{pmatrix} .$$

and also

$$(Q - P')P' = \frac{(a_2 - 1)}{(\gamma_3 - \gamma_2)\gamma_2} \begin{pmatrix} 0 & a_{11} \\ 0 & a_{21} \end{pmatrix}$$

$$P'(Q - P') = \frac{a_{21}/\gamma_2}{(\gamma_3 - \gamma_2)} \begin{pmatrix} a_2 & -a_2/\gamma_2 \\ -\gamma_2 & 1 \end{pmatrix}$$

The difference in lower components at time u is thus

$$\text{I:} \quad \frac{\delta}{(\gamma_3 - \gamma_2)} \left[a_{21}(c_2 - c_3/\gamma_2) + \frac{(a_2+1)}{\gamma_2} a_{21}c_3u + o(u) \right]$$

$$\text{II:} \quad \frac{\Delta}{(\gamma_3 - \gamma_2)} \left[a_{21}(c_2 - c_3/\gamma_2) - a_{21}(c_2 - c_3/\gamma_2)u + o(u) \right] .$$

The difference in upper components is

$$\text{I:} \quad \frac{\delta}{(\gamma_3 - \gamma_2)} \left[a_{11}(c_2 - c_3/\gamma_2) + \frac{(a_2+1)}{\gamma_2} a_{11}c_3u + o(u) \right].$$

$$\text{II:} \quad \frac{\Delta}{(\gamma_3 - \gamma_2)} \left[a_{11}(c_2 - c_3/\gamma_2) + \frac{a_{21}a_2}{\gamma_2} (c_2 - c_3/\gamma_2)u + o(u) \right].$$

We choose Δ so that the policy of type II gives the same value for $x_3(u)$ as that given by the policy of type I. Thus

$$\Delta = \frac{\delta\left[(c_2-c_3/\gamma_2) + \frac{(a_2+1)}{\gamma_2} c_3 u + o(u)\right]}{(c_2-c_3/\gamma_2)\ (1-u)} .$$

Now we prove that the quantity $x_2(u)$ for II exceeds $x_2(u)$ for I by a positive amount. Since $(\gamma_3-\gamma_2) < 0$, $(c_2-c_3/\gamma_2) > 0$, and $(1-u) > 0$, this is equivalent to the statement

$$a_{11}\left[(c_2-c_3/\gamma_2) + \frac{(a_2+1)}{\gamma_2} c_3 u + o(u)\right](1-u)$$

$$> \left[(c_2-c_3/\gamma_2) + \frac{(a_2+1)}{\gamma_2} c_3 u + o(u)\right](a_{11} + \frac{a_{21}}{\gamma_2} a_2 u + o(u)).$$

The quantity in brackets is positive because u is small. Hence this statement is equivalent to

$$a_{11}(1-u) > (a_{11} + a_{21}\frac{a_2}{\gamma_2} u + o(u)).$$

But $a_{21}\frac{a_2}{\gamma_2} + a_{11} = -a_2\gamma_3(b_3+1) + a_2\gamma_3 + \gamma_2 = -(a_2 b_3\gamma_3 - \gamma_2) - \lambda < 0$, so that the above inequality holds for u and δ/u sufficiently small. We see that with the above choice of Δ the policy of type II is superior.

§7. Q Policy not Preceded by R.

Now, we prove that an optimal policy cannot have the form: Follow R for $0 < t < \delta$; follow Q for $\delta < t < T - T^*$; follow P' for $T - T^* < t < T$. This will complete the proof that Q is followed

in the whole region $1/\gamma_2 < r < 1/\gamma_3$, $T > T^*$. We show that a policy of the above form is not optimal by comparing it with a policy of the form: Follow Q for $0 < t < T - T^*$; follow P' for $T - T^* < t < T$.

As before we let P' and Q denote the matrices for which $dx/dt = P'x$ and $dx/dt = Qx$ when P' and Q respectively are followed. Similarly we let

$$R = \begin{pmatrix} -1 & 0 \\ b_3\gamma_3 & 0 \end{pmatrix}.$$

For the policies described above

$$\text{I:}\quad x(T) = e^{P'T^*} e^{Q(T-T^*)} e^{-Q\delta} \cdot e^{R\delta} \cdot c$$

$$= e^{P'T^*} e^{Q(T-T^*)} (I + (R-Q)\delta + o(\delta)) \cdot c$$

$$\text{II:}\quad x(T) = e^{P'T^*} e^{Q(T-T^*)} \cdot c.$$

Now

$$(R - Q) = \frac{1}{(\gamma_3-\gamma_2)} \begin{pmatrix} -\gamma_3(a_2+1) & (a_2+1) \\ \gamma_3(b_3\gamma_3+\gamma_2) & -(b_3\gamma_3+\gamma_2) \end{pmatrix}.$$

Let $e^{P'T^*} e^{Q(T-T^*)} = \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{pmatrix}$. Then for the two policies we have respectively

$$\text{I:}\quad x_2(T) = \alpha_{11}c_2 + \alpha_{12}c_3 + \frac{\delta(c_3-\gamma_3 c_2)}{(\gamma_3-\gamma_2)}\left[\alpha_{11}(a_2+1)-\alpha_{12}(b_3\gamma_3+\gamma_2)\right] + o(\delta),$$

$$\text{II:}\quad x_2(T) = \alpha_{11}c_2 + \alpha_{12}c_3.$$

Since $(\gamma_3-\gamma_2) < 0$, $(c_3-\gamma_3 c_2) > 0$, it will be sufficient to prove that

$$\alpha_{11}(a_2+1) - \alpha_{12}(b_3\gamma_3+\gamma_2) > 0.$$

Note that α_{11} and α_{12}, although they depend on T, do not depend on r. Consequently, if we succeed in verifying the inequality for a particular r for which $1/\gamma_2 < r < 1/\gamma_3$, we will have proved the desired result. There is a value r^* in this interval with the property that if Q is followed and $c_2/c_3 = r^*$ then $x_2(t)/x_3(t) = r^*$ as long as Q is followed. We postpone the proof of the existence of such an r^* for the moment and continue the above proof, with the additional assumption that $c_2/c_3 = r^*$. In this particular case

$$e^{Q(T-T^*)} = \begin{pmatrix} \mu(T) & 0 \\ 0 & \mu(T) \end{pmatrix}$$

where $\mu(T) > 0$ depends only on T. Since

$$\begin{aligned} x_2(T) &= x_2(T-T^*) + \frac{a_2}{\gamma_2} x_3(T-T^*)(1-e^{-T^*}) \\ &= x_2(T-T^*) + x_3(T-T^*)\,\frac{(a_2\gamma_3 + \gamma_2)}{\gamma_2\gamma_3(b_3 + 1)} \end{aligned}$$

we have

$$e^{P'T^*} = \begin{pmatrix} 1 & \dfrac{(a_2\gamma_3 + \gamma_2)}{\gamma_2\,\gamma_3(b_3 + 1)} \\ * & * \end{pmatrix}$$

where the asterisks represent numbers in which we are not interested.

Consequently

$$e^{P'T^*}\, e^{Q(T-T^*)} = \begin{pmatrix} \mu(T) & \mu(T)\,\dfrac{(a_2\gamma_3+\gamma_2)}{\gamma_2\gamma_3(b_3+1)} \\ * & * \end{pmatrix}.$$

We want to prove

$$\mu(T)\left[(a_2+1) - (b_3\gamma_3+\gamma_2)\,\frac{(a_2\gamma_3+\gamma_2)}{\gamma_2\gamma_3(b_3+1)}\right] > 0.$$

or

$$\gamma_2\gamma_3(a_2+1)(b_3+1) - (b_3\gamma_3+\gamma_2)(a_2\gamma_3+\gamma_2) > 0$$

or

$$(\gamma_2-\gamma_3)(a_2b_3\gamma_3-\gamma_2) = (\gamma_2-\gamma_3)\lambda > 0$$

which holds since $\gamma_2 > \gamma_3$ and $\lambda > 0$. Except for the proof of the existence of r^* this completes the proof.

§8. Effect on r of Q Policy.

Let us investigate the effect on the ratio $r = x_2/x_3$ of following Q. Since this will be important also in the case $\gamma_3/\gamma_2 > 1$, we consider both cases at the same time. Since $x_2 = rx_3$, we have

$$\frac{dx_2}{dt} = \frac{dr}{dt} + r\,\frac{dx_3}{dt}.$$

Now, if we let

$$Q = \frac{1}{(\gamma_3-\gamma_2)}\begin{pmatrix} a_2\gamma_3+\gamma_2 & -(a_2+1) \\ -\gamma_2\gamma_3(b_3+1) & b_3\gamma_3+\gamma_2 \end{pmatrix} = \frac{1}{(\gamma_3-\gamma_2)}\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix},$$

we have

$$a_{11}x_2 + a_{12}x_3 = \frac{dr}{dt}x_3(\gamma_3-\gamma_2) + ra_{21}x_2 + ra_{22}x_3$$

or

$$-(\gamma_3-\gamma_2)\,\frac{dr}{dt} = a_{21}r^2 + (a_{22}-a_{11})r - a_{12}.$$

The zeros of the right side are

$$\frac{a_{11}-a_{22}}{2a_{21}} \pm \frac{1}{2}\sqrt{\left(\frac{a_{11}-a_{22}}{a_{21}}\right)^2 + 4\,\frac{a_{12}}{a_{21}}}\quad .$$

There is one positive zero and one negative zero. We define r^* to be the positive zero

$$r^* = \frac{a_{11}-a_{22}}{2a_{21}} + \frac{1}{2}\sqrt{\left(\frac{a_{11}-a_{22}}{a_{21}}\right)^2 + 4\,\frac{a_{12}}{a_{21}}}$$

$$= \frac{(b_3-a_2)}{2\gamma_2(b_3+1)} + \frac{1}{2}\sqrt{\left(\frac{b_3-a_2}{\gamma_2(b_3+1)}\right)^2 + \frac{4(a_2+1)}{\gamma_2\gamma_3(b_3+1)}}\quad .$$

Also r^* is between $1/\gamma_3$ and $1/\gamma_2$. The easiest way to verify this is to compute the value of the quadratic on the above right side for $r = 1/\gamma_2$ and $r = 1/\gamma_3$. The values are

$$\frac{-(a_2+1)(\gamma_3-\gamma_2)}{\gamma_2} \quad\text{and}\quad \frac{(b_3+1)(\gamma_3-\gamma_2)}{\gamma_3}$$

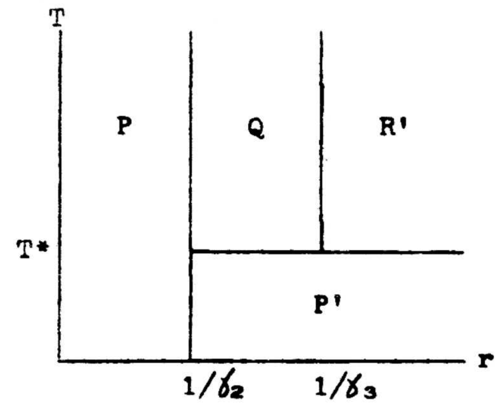
which differ in sign.

When the Q policy is followed and $r = r^*$, r does not change. It is also easy to see from the differential equation above that for $\gamma_3/\gamma_2 < 1$, r increases when $r < r^*$ and r decreases when $r > r^*$. On the other hand, for $\gamma_3/\gamma_2 > 1$, r increases when $r > r^*$ and decreases when $r < r^*$.

§9. Remaining Regions.

We finally conclude that Q is followed in the entire region $1/\gamma_2 < r < 1/\gamma_3$, $T > T^*$. It follows that in this same region $\gamma_3/\gamma_2 < B/A < 1$. Since B/A increases monotonely as r increases, we can conclude that $B/A < 1$ for $r < 1/\gamma_2$, $T > T^*$ and $B/A > \gamma_3/\gamma_2$ for $r > 1/\gamma_3$, $T > T^*$. From this we conclude P is followed in the region $r < 1/\gamma_2$, $T > T^*$ and R' in the region $r > 1/\gamma_3$, $T > T^*$.

The solution has now been found everywhere except in the region $r < 1/\gamma_2$, $T \leq T^*$. A natural conjecture is that P is followed in this region also. This can be checked directly, but it also follows from the monotonicity of B/A. We know that for $T < T^*$, $r > 1/\gamma_2$, the function $B/A < \gamma_3/\gamma_2 < 1$. Hence in the region $T < T^*$, $r < 1/\gamma_2$, $B/A < 1$ and P is followed.



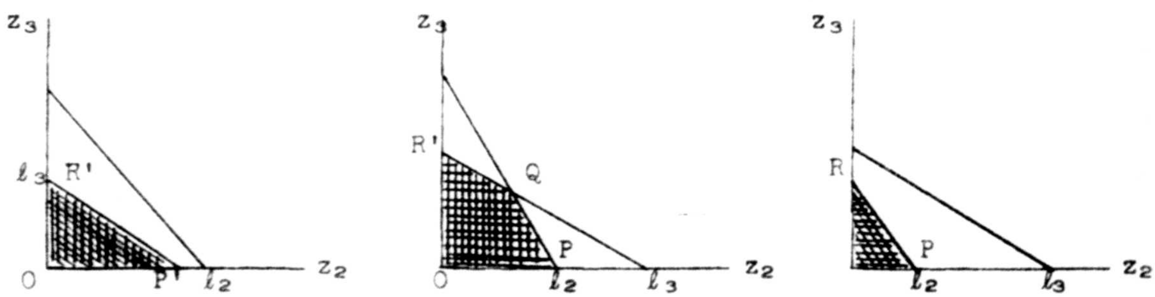
§10. <u>The Case</u> $\gamma_3/\gamma_2 > 1$.

Now we consider the case in which $\gamma_3/\gamma_2 > 1$. Again there are three different cases depending on the position of the lines

$$\ell_2: \quad z_2 + z_3 = x_2$$
$$\ell_3: \quad \gamma_2 z_2 + \gamma_3 z_3 = x_3.$$

We use the same letters to denote the vertices as in the case $\gamma_3/\gamma_2 < 1$.

$r > 1/\gamma_2$ $\qquad$ $1/\gamma_3 < r < 1/\gamma_2$ $\qquad$ $r < 1/\gamma_3$



The optimal policies are determined by the ratio as follows:

$r > 1/\gamma_2$	$1/\gamma_3 < r < 1/\gamma_2$	$r < 1/\gamma_3$
$B > \frac{\gamma_3}{\gamma_2} A$: R'	$B > \frac{\gamma_3}{\gamma_2} A$: R'	$B > A$: R
$B < \frac{\gamma_3}{\gamma_2} A$: P'	$A < B < \frac{\gamma_3}{\gamma_2} A$: Q	$B < A$: P.
	$B < A$: P	

The results obtained in the previous case shows that for $r > 1/\gamma_2$, $T < T^*$, the optimal policy is to follow P'. There is another important point which can be calculated using those results, the point on the line $r = 1/\gamma_2$ where $B = A$. Let T_0 be that point if it exists and otherwise let $T_0 = \infty$. We have

$$a_2 e^{-T_0} = \frac{b_3\gamma_3 a_2}{\gamma_2}(1 - e^{-T_0}) - 1$$

$$e^{-T_0} = \frac{a_2 b_3\gamma_3 - \gamma_2}{a_2 b_3\gamma_3 + a_2\gamma_2} = \frac{\lambda}{a_2(b_3\gamma_3 + \gamma_2)}$$

A necessary and sufficient condition that $T_0 \neq \infty$ is $\lambda > 0$. Also, when $\lambda > 0$, $T_0 < T^*$ because $\gamma_3 > \gamma_2$.

Because of the monotonicity of B/A as a function of r, $B < A$ in the region $T < T_0$, $r < 1/\gamma_2$; and hence the policy P is followed there. Our next aim is to determine the location of the region where Q is followed.

§11. Q Policy Does Not Precede P.

First, we establish that in an optimal policy it does not

happen that first Q is followed and then P. The method is the same as that used previously. Consider any policy which consists in following Q for $0 < t < \delta$ and P for $\delta < t < u$. We prove that for δ and δ/u sufficiently small there is a superior policy having the form P for $0 < t < u - \Delta$ and Q for $u - \Delta < t < u$ where Δ is a positive number.

Let

$$P = \begin{pmatrix} a_2 & 0 \\ -\gamma_2 & 0 \end{pmatrix},$$

and Q be the matrix for which $dx/dt = Qx$ when Q is followed. For the policies described above we have

$$\text{I:} \quad x(u) = e^{P(u-\delta)}\, e^{Q\delta} \cdot c = e^{Pu}(I + (Q-P)\delta + o(\delta)) \cdot c$$
$$\text{II:} \quad x(u) = e^{Q\Delta} e^{P(u-\Delta)} \cdot c = (I + (Q-P)\Delta + o(\Delta))e^{Pu} \cdot c$$

For δ/u, Δ/u, $u \longrightarrow 0$, the difference between the result at time u obtained by following one of these policies and $e^{Pu} \cdot c$ is

$$\text{I:} \quad \delta \left[(Q-P) + P(Q-P)u + o(u) \right] \cdot c$$
$$\text{II:} \quad \Delta \left[(Q-P) + (Q-P)Pu + o(u) \right] \cdot c$$

Setting $a_{12} = -(a_2+1)$, $a_{22} = b_3\gamma_3 + \gamma_2$, we find

$$(Q-P) = \frac{1}{(\gamma_3-\gamma_2)} \begin{pmatrix} -\gamma_2 a_{12} & a_{12} \\ -\gamma_2 a_{22} & a_{22} \end{pmatrix}$$

$$P(Q-P) = \frac{a_{12}}{(\gamma_3-\gamma_2)} \begin{pmatrix} -\gamma_2 a_2 & a_2 \\ \gamma_2^2 & -\gamma_2 \end{pmatrix}$$

$$(Q-P)P = \frac{-(a_2+1)\gamma_2}{(\gamma_3-\gamma_2)} \begin{pmatrix} a_{12} & 0 \\ a_{22} & 0 \end{pmatrix}$$

The difference in lower components at time u is therefore

$$\text{I:} \qquad \frac{\delta}{(\gamma_3-\gamma_2)} \Big[a_{22}(c_3-c_2\gamma_2) - a_{12}\gamma_2(c_3-c_2\gamma_2)u + o(u) \Big]$$

$$\text{II:} \qquad \frac{\Delta}{(\gamma_3-\gamma_2)} \Big[a_{22}(c_3-c_2\gamma_2) - (a_2+1)\gamma_2 a_{22} c_2 u + o(u) \Big].$$

The difference in upper components is

$$\text{I:} \qquad \frac{\delta}{(\gamma_3-\gamma_2)} \Big[a_{12}(c_3-c_2\gamma_2) + a_{12}a_2(c_3-c_2\gamma_2)u + o(u) \Big]$$

$$\text{II:} \qquad \frac{\Delta}{(\gamma_3-\gamma_2)} \Big[a_{12}(c_3-c_2\gamma_2) - a_{12}(a_2+1)\gamma_2 c_2 u + o(u) \Big].$$

We choose Δ so that the policy II gives the same value for $x_3(u)$ as that given by policy I. We have

$$\Delta = \frac{\delta(c_3-c_2\gamma_2)\,\big[a_{22}-a_{12}\gamma_2 u + o(u)\big]}{a_{22}\,\big[(c_3-c_2\gamma_2) - (a_2+1)\gamma_2 c_2 u\big]}$$

Now we prove the quantity $x_2(u)$ for II exceeds that for I by a positive amount. Since $(\gamma_3-\gamma_2) > 0$, $(c_3-c_2\gamma_2) > 0$, $a_{22} > 0$, this is equivalent to the statement

$$a_{12}\left[(c_3-c_2\gamma_2)-(a_2+1)\gamma_2c_2u+o(u)\right]\left[a_{22}-a_{12}\gamma_2u+o(u)\right]$$

$$>a_{12}\left[1+a_2u+o(u)\right]a_{22}\left[(c_3-c_2\gamma_2)-(a_2+1)\gamma_2c_2u\right].$$

This is equivalent to

$$\left[a_{22}-a_{12}\gamma_2u+o(u)\right]<a_{22}\left[1+a_2u+o(u)\right].$$

But $a_{22}a_2+a_{12}\gamma_2=a_2(b_3\gamma_3+\gamma_2)-\gamma_2(a_2+1)=\lambda>0$. Consequently, the above inequality holds for u and δ/u sufficiently small.

§12. Q Policy Does Not Precede R'.

Next, we establish that in an optimal policy it does not happen that first a Q policy is followed and then R'. Consider any policy which consists in following Q for $0<t<\delta$ and R' for $\delta<t<u$. We prove that for u and δ/u sufficiently small there is a superior policy having the form R' for $0\leq t<u-\Delta$ and Q for $u-\Delta<t<u$, where Δ is a positive number. Let $R'=\begin{pmatrix}0 & -1/\gamma_3\\ 0 & b_3\end{pmatrix}$. For the policies described above we have

$$\text{I:}\quad x(u)=e^{R'(u-\delta)}e^{Q\delta}\cdot c=e^{R'u}(I+(Q-R')\delta+o(\delta))\cdot c$$

$$\text{II:}\quad x(u)=e^{Q\Delta}e^{R'(u-\Delta)}\cdot c=(I+(Q-R')\Delta+o(\Delta))e^{R'u}\cdot c$$

For δ/u, Δ/u, $u\longrightarrow 0$ the difference between the result at time u obtained by following one of these policies and $e^{R'u}\cdot c$ is

$$\text{I:}\quad \delta\left[(Q-R')+R'(Q-R')u+o(u)\right]\cdot c$$

$$\text{II:}\quad \Delta\left[(Q-R')+(Q-R')R'u+o(u)\right]\cdot c.$$

Setting $a_{11}=a_2\gamma_3+\gamma_2$, $a_{21}=-\gamma_2\gamma_3(b_3+1)$, we find

$$(Q-R') = \frac{1}{(\gamma_3-\gamma_2)} \begin{pmatrix} a_{11} & -a_{11}/\gamma_3 \\ a_{21} & -a_{21}/\gamma_3 \end{pmatrix},$$

$$R'(Q-R') = \frac{a_{21}}{(\gamma_3-\gamma_2)} \begin{pmatrix} -1/\gamma_3 & 1/\gamma_3^2 \\ b_3 & -b_3/\gamma_3 \end{pmatrix}$$

$$(Q-R')R' = \frac{-(b_3+1)}{\gamma_3(\gamma_3-\gamma_2)} \begin{pmatrix} 0 & a_{11} \\ 0 & a_{21} \end{pmatrix}$$

The difference in lower components at time u is therefore

$$\text{I:} \quad \frac{\delta}{(\gamma_3-\gamma_2)} \left[a_{21}(c_2-c_3/\gamma_3) + a_{21}b_3(c_2-c_3/\gamma_3)u + o(u) \right],$$

$$\text{II:} \quad \frac{\Delta}{(\gamma_3-\gamma_2)} \left[a_{21}(c_2-c_3/\gamma_3) - a_{21}\frac{(b_3+1)}{\gamma_3} c_3u + o(u) \right].$$

The difference in upper components is

$$\text{I:} \quad \frac{\delta}{(\gamma_3-\gamma_2)} \left[a_{11}(c_2-c_3/\gamma_3) - \frac{a_{21}}{\gamma_3}(c_2-c_3/\gamma_3)u + o(u) \right]$$

$$\text{II:} \quad \frac{\Delta}{(\gamma_3-\gamma_2)} \left[a_{11}(c_2-c_3/\gamma_3) - a_{11}\frac{(b_3+1)}{\gamma_3} c_3u + o(u) \right].$$

We choose Δ so that the policy II gives the same value for $x_3(u)$ as that given by policy I. We have

$$\Delta = \frac{\delta(c_2-c_3/\gamma_3)\,\left[1+b_3u+o(u)\right]}{(c_2-c_3/\gamma_3) - \frac{(b_3+1)}{\gamma_3}c_3u}$$

We prove that with this choice of Δ the quantity $x_2(u)$ for II exceeds $x_2(u)$ for I by a positive amount. Since $(\gamma_3-\gamma_2) > 0$,

$(c_3-c_3/\gamma_3) > 0$, this is equivalent to the statement

$$a_{11}\left[1 + b_3u + o(u)\right]\left[(c_2-c_3/\gamma_3) - \frac{(b_3+1)}{\gamma_3}c_3u\right]$$

$$> \left[a_{11} - \frac{a_{21}}{\gamma_3}u + o(u)\right]\left[(c_2-c_3/\gamma_3) - \frac{(b_3+1)}{\gamma_3}c_3u\right],$$

which is equivalent to

$$a_{11}(1 + b_3u + o(u)) > a_{11} - \frac{a_{21}}{\gamma_3}u.$$

But

$$a_{11}b_3 + \frac{a_{21}}{\gamma_3} = a_2b_3\gamma_3 + b_3\gamma_2 - \gamma_2(b_3+1) = \lambda > 0.$$

Consequently the above inequality holds for u and δ/u sufficiently small.

§13. Location of Q Region.

Moreover, since a Q policy is not followed for $T < T_o$, we conclude that a Q policy is always succeeded by the P' policy. Now we prove that the region where the Q policy is followed does not contain any points (r',T') where $r' < r^*$. Suppose the contrary. Then, since, as in §8, r decreases when the Q policy is followed for $r < r^*$, there are points (r,T) in the Q region for which $r < r^*$, $T < T_o$. But this contradicts a result which states that all such points are in the region where P is followed.

Consider the Riccati differential equation in §8 which represents the dependence of r on t when the Q policy is followed.

For $r^* < r_o \leq 1/\gamma_2$ let $\phi(r_o)$ be the time required to reach $r = 1/\gamma_2$ beginning from r_o. From the differential equation we see that $\phi(r)$ is a monotone decreasing function of r and $\phi(r) \longrightarrow +\infty$ as $r \longrightarrow r^*$. Also $\phi(1/\gamma_2) = 0$.

The result that a Q policy is always succeeded by a P' policy permits us to reach some conclusions about the location of the Q region. In fact, if (r,T) is a point in the Q region then

$$T_o \leq T - \phi(r) \leq T^*,$$

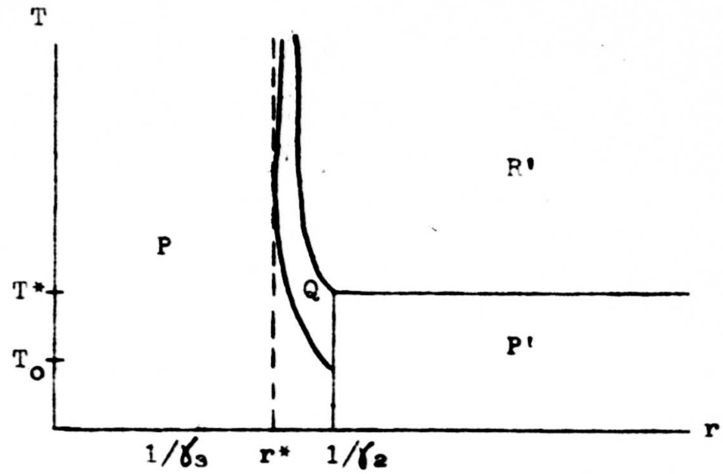
and consequently

$$T_o + \phi(r) \leq T \leq T^* + \phi(r).$$

There are points (r,T) in the Q region with arbitrarily large T. To verify this we consider the continuous curve on which $B/A = 1$ which passes through the point $(1/\gamma_2, T_o)$. Near the point $(1/\gamma_2, T_o)$ it forms one of the boundaries of the Q region, and it will continue to do so as T increases since by the above results it can never leave the region $r^* < r \leq 1/\gamma_2$.

§14. Remaining Regions.

In the Q region $1 \leq B/A \leq \gamma_3/\gamma_2$. By the monotonicity of B/A as a function of r, we infer that to the right of the Q region for $T > T^*$ the optimal policy is R', while to the left of the Q region the optimal policy is P.



This completes the finding of the solution except for one possibility. What policy is followed on the boundaries of the regions? It will make no difference if a solution trajectory crosses the boundaries, but if the solution trajectory runs along one of the boundary curves we cannot even be sure that an optimal policy is obtained by choosing one of the vertices P, P', Q, R or R'. We can eliminate this possibility by showing that to follow the P policy on the P-Q boundary would drive us into the Q region, and that similarly to follow the R' policy on the Q-R' boundary would drive us into the Q region. For we already know that the Q policy keeps us in the Q region; hence it would follow that any convex linear combination with one of the policies P or R' would drive us into the Q region.

First, let us consider the result of following P on the P-Q boundary. Looking at the differential equations (2.1), we see that following P rather than Q increases x_2 more, increases x_3 less, and consequently increases r more. Hence the solution trajectory must go into the Q region. A similar argument gives the same result on the Q-R' boundary.

Finally, we note that there is also the intermediate case $\gamma_3/\gamma_2 = 1$, which forms the natural connection between the two cases already discussed. As should be expected the policies followed are:

For $r < 1/\gamma_2 = 1/\gamma_3$: P

For $r > 1/\gamma_2$, $T < T^*$: P'

For $r > 1/\gamma_2$, $T > T^*$: R'

In this case the boundary between the P and R' regions is of interest. To follow P would drive us into the R' region, while to follow R' would drive us into the P region. But there is an intermediate policy which keeps us on the boundary. The policy is given by

$$z_2 = \frac{(b_3+1)x_2}{(a_2+b_3+2)}, \quad z_3 = \frac{(a_2+1)x_2}{(a_2+b_3+2)}$$

It is this policy which is followed. This illustrates that care must be used in assuming that an optimal policy consists only of vertex choices.

BIBLIOGRAPHY

1. Bellman, R., "Bottleneck Problems and Dynamic Programming," Proc. Nat. Acad. Sci., 39 (Sept. 1953), pp. 947-51.

2. ————, "Bottleneck Problems, Functional Equations, and Dynamic Programming," RAND Paper No. P-483 (submitted to Econometrica).

AFBMD Library Charge-Out Card

Document No: 55-2437 Copy No: 1 O/A Report No: P 490

Title: SOME APPLICATIONS OF THE THEORY OF DYNAMIC PROGRAMMING A REVIEW

Class: (U)

Orig. Ag'y: RAND

Author: RICHARD BELLMAN Doc. Date: 14 May 54 Contract No:

Borrower's Name	Date Out	Date In	Borrower's Signature
PATTI LEE, STL	9/22/59	10/27	Patti Lee